Statement of Purpose

Nirlipta Pande
nirliptapande@gmail.com

In Jewish folklore, a golem is a clay robot brought to life by inscribing truth on its brows. It lacks free will and does exactly what is told, which is good because it is extremely powerful and can accomplish far more than its creators can. Such are statistical methods like t-tests, set of design principles and constraints to crunch numbers and perform impressive calculations. I've created golems trained for diverse tasks, from building a decentralized voting system to developing causal models for global ignitions. Such research experiences helped me realize my long-term goal of developing responsible and explainable models to help us simulate the real world. As a pre-doctoral researcher at Google, I intend to work on developing causal representation for real-world modelling to ease inclusion of System 2 reasoning in existing models.

Good forecasts can seem magical but pattern detection-based speculations sometimes lead to outcomes like the 2008 subprime mortgage crisis. I wanted to understand such complex variable dynamics hence I chose to major in computer science and economics at BITS Pilani . At ADAPT lab, I worked on crop yield forecasting by combining optical and radar datasets. We used rule mining to uncover proxies for variables like soil moisture followed by temporal optimization. We took inspiration from You et al. for dimensionality reduction and devising a Deep Gaussian Process framework. Since Gaussian processes are computationally intensive, I developed baseline model by combining autoregressive models and kernels to extract the linear and non-linear patterns, respectively. This project exposed me to practical deep learning and time series which I built on during my thesis under Prof. Atzberger on nowcasting soil moisture. In such short forecasting windows, robustness and scalability are hard to optimize. While Kalman filters were the obvious choice owing to their accuracy, we needed faster alternatives for near real-time models. I fitted non-linear edges to smaller data sections leading to of the original runtime with minimal loss in accuracy.

All this while, I started exploring causal inference due to the lack of modelling inter-variable relationships. I approached Prof. Dorigo to collaborate on developing causal models for global ignitions where we identified emergent relationships between socioeconomic factors, burned areas, and climatological factors. A causal perspective highlights that the tropical region is prone to fires due to the climate and landcover and not just the anthropogenic interventions and helps separate their causal effect. In this multi-dimensional dataset, developing regional structural models followed by confirmatory analysis over each pixel is an efficient way which we presented at the EGU General Assembly 2023. I was also involved in profiling and rewriting the codebase for faster implementation leading to our paper on anthropogenic effects analyzing ignitions.

I believe in putting science before statistics, which is very important for research transferable into the real-world. For instance, by building LLMs with a causal structure, we will be able to detect and remove data and model bias, while LLMs can identify underlying causal structures in a defined system. Such research will help me lay strong foundations for my research goals of pursuing a PhD with a direct societal impact.

Thank you for considering my application.